A (72,36;15) Box Code

G. Solomon[1]
Communications Systems Research Section

A (72,36;15) box code is constructed as a 9 × 8 matrix whose columns add to form an extended BCH-Hamming (8,4;4) code and whose rows sum to odd or even parity. The newly constructed code, due to its matrix form, is easily decodable for all seven-error and many eight-error patterns. The code comes from a slight modification in the parity (eighth) dimension of the Reed-Solomon (8,4;5) code over GF*(512). Error correction uses the row sum parity information to detect errors, which then become erasures in a Reed-Solomon correction algorithm.*

I. The Code Construction

The first 27 dimensions of the codes constructed basically constitute a (63,27;16) code represented as a 9×7 matrix. This arises from a Reed-Solomon (RS) cyclic (7,3;5) code over $GF(512)$. The last nine dimensions are constructed by modifying the construction of the extended RS (8,4;5) code over $GF(512)$. For eight of the nine dimensions, this is exactly the extended Reed-Solomon code. For the ninth dimension, the encoding algorithm is modified. Encoding is direct and systematic. Decoding the code uses error/erasure techniques as discussed in [1].

II. A Modified Reed-Solomon (8,4;5) Code Over $GF(512)$

Recall first that the extended RS (8,4;5) code over $GF(8)$ represented in binary form in a normal basis is isomorphic to the extended Golay (24,12;8) code. However, the modified extended RS (8,4;5) code over $GF(64)$ represented in binary form is a self-dual (48,24;12) code. These were shown in [1]. In both these codes, the decoding must sometimes go through eight such trials corresponding to an ambiguity of elements in $GF(8)$. The decoding procedure here will exhibit the same ambiguity for one special case of seven errors that appear as erasures.

Using techniques similar to those in [1], if one starts with an RS (8,4;5) code over $GF(512)$, and represents the code in binary using a particular normal basis with the special property defined below, one can generate a code of length 72 and dimension 36 with even weights that are multiples of 4 and odd weights of form $4m - 1$.

The binary representation of the usual RS (8,4;5) code over $GF(512)$, yields nine (8,7;2) codewords whose decomposition into two cyclic code components and a constant component looks, respectively, like (9,7;3) and (9,2;8) RS or maximal-distance-separable (MDS) codes over $GF(8)$ and a binary (9,9;1) code. However, the code here is constructed by modifying the extended coding rule for the parity symbol.

In particular, let γ be a root of the polynomial $f(x) = x^9 + x^8 + x^6 + x^5 + x^4 + x + 1$, where γ is a primitive generator of the 511 roots of unity. Represent the elements of $GF(512)$ in the normal representation using

[1] Consultant.

the roots of $f(x)$. The roots are γ^j, where $j \in J$ for $J = \{1, 2, 4, 8, 16, 32, 64, 128, 256\}$.

Note that for this particular choice of $f(x)$, one has

$$\text{Tr } \gamma^j = 1, \quad j \in J, \quad J = \{1, 2, 4, 8, 16, 32, 64, 128, 256\}$$

$$\text{Tr } \gamma^i \gamma^k = 0, \quad i \neq k, \quad i, k \in J$$

Let β be a root of the polynomial $g(x) = x^3 + x^2 + 1$. Here β is an element of $GF(8)$, a subfield of $GF(512)$, and $\beta = \gamma^{73}$.

A. Encoding

Now use the recursion or check polynomial $h(x) = \prod_{i=0}^{3}(x+\beta^i)$ to generate an RS (7,4;4) code over $GF(512)$. This means that the initial shift register contains four elements in $GF(512)$ expressed as coefficients in the normal representation above. The cyclic portion of the code is of length seven, and the overall parity symbol, the eighth dimension, is defined differently. Representing the binary code as components Tr $P(x)\gamma^i$, where $i = 1, 2, 4, 8, 16, 32, 64, 128, 256$, extends the codes to the eighth coordinates by the rules; the binary value at the row indexed by the ith coordinate is given by Tr $C_0\gamma^i$ + Tr $\sum_{j \in J} C_0\gamma^i$.

Thus, for the constant term C_0 with Tr $C_0 = 0$, this symbol behaves like the normal parity symbol, which is a sum over the values of the cyclic code coordinates.

The general Mattson-Solomon (MS) polynomial of a codeword $\mathbf{a}$ is $P_{\mathbf{a}}(x) = C_0 + C_1x + C_2x^2 + C_3x^3$, where $C_i \in GF(512)$ for $0 \leq i \leq 3$ and $x \in GF(8)$. Encode the codeword in its cyclic portion. The extended codeword $\mathbf{a}$ expressed in terms of the MS polynomial is

$$\mathbf{a} = \left(P_{\mathbf{a}}(\beta^i), \quad 0 \leq i \leq 6, \quad P_{\mathbf{a}}(0)\right)$$

Writing the codewords in binary and using the normal basis γ^j for $j \in J$ above, there are nine binary codewords of length eight

$$\text{Tr } P(x)\gamma^j, \quad j = 1, 2, 4, 8, 16, 32, 64, 128, 256$$

where Tr a denotes the value in $GF(2)$ given by the trace of an element $a \in GF(512)$

$$\text{Tr } a = a + a^2 + a^4 + a^8 + a^{16} + a^{32} + a^{64} + a^{128} + a^{256}$$

Consider one of the nine binary words in its Mattson-Solomon setting,

$$\begin{aligned}\text{Tr } P_{\mathbf{a}}(x)\gamma^j &= \text{Tr } (C_0 + C_1x + C_2x^2 + C_3x^3)\gamma^j \\ &= \text{Tr } C_0\gamma^j + \text{Tr }'[(C_1x + C_2x^2 + C_3x^3)\gamma^j] \\ &\quad + \text{Tr }'[((C_1x + C_2x^2 + C_3x^3)\gamma^j)^8 \\ &\quad + ((C_1x + C_2x^2 + C_3x^3)\gamma^j)^{64}]\end{aligned}$$

$$\text{Tr }'a = a + a^2 + a^4, \quad a \in GF(8)$$

Set $C_0 = 0$ temporarily, as this does not affect the arguments to follow. Then,

$$\begin{aligned}\text{Tr } P(x)\gamma^j = \text{Tr }'&(C_1\gamma^j + (C_1\gamma^j)^8 + (C_1\gamma^j)^{64} \\ &+ (C_2\gamma^j)^{256} + (C_2\gamma^{256j})^8 + (C_2\gamma^{256j})^{64})x \\ &+ \text{Tr }'((C_3\gamma^j)^2 + (C_3\gamma^j)^{16} + (C_3\gamma^j)^{128})x^6\end{aligned}$$

Lemma. The coefficient of x is a (9,6:4) code over $GF(8)$. The coefficient of x^3, and consequently of x^6, is a (9,3;7) code over $GF(8)$. The code is indexed by the values of γ^j, where $j \in J = \{1, 2, 4, 8, 16, 32, 64, 128, 256\}$.

Proof. The set γ^j, where $j \in J = \{1, 2, 4, 8, 16, 32, 64, 128, 256\}$ can only take zero values one less than the number of terms in the coefficients of x and x^6. An argument that clarifies this follows. Consider the coefficient of x^3. This is a polynomial of degree 64 for which, if γ is a solution for a value C_3, then $A\gamma$ is also a solution for all $A \in GF(8)$. Thus, there can only be at most two values of γ that make the coefficient of x^3 equal to zero and consequently the coefficient of x^6 equal to zero. A similar examination of the degrees in the coefficient of x will yield the above result. The term Tr $C_0\gamma^j$ in the code's expression when Tr $C_0 = 0$, i.e., the constant terms, forms a binary (9,8;2) code.

Theorem. The RS code determined by codewords with MS polynomials $P_{\mathbf{a}}(x)$ and Tr $C_0 = 0$ forms a binary (72,35;16) code with weights that are multiples of 4.

Proof. The multiple-of-4 property of the weights using the Solomon-McEliece Γ_2 Formula follows:

$$\text{Tr } P(x)\gamma = \text{Tr } ((C_1\gamma + (C_2\gamma)^4 + (C_3\gamma^2)x^6)$$

where Tr is defined in $GF(64)$.

Now

$$\begin{aligned}\Gamma_2(\text{Tr } P(x)\gamma) = \text{Tr } [&C_1C_3^2\gamma^3 + C_1^8C_3^2\gamma^{10} + C_1^{64}C_3^2\gamma^{66} \\ &+ C_1C_3^{16}\gamma^{17} + C_1^8C_3^{16}\gamma^{24} + C_1^{64}C_3^{16}\gamma^{80} \\ &+ C_1C_3^{128}\gamma^{129} + C_1^8C_3^{128}\gamma^{136} + C_1^{64}C_3^{128}\gamma^{192} \\ &+ C_2^{256}C_3^2\gamma^{258} + C_2^{256}C_3^{16}\gamma^{272} + C_2^{256}C_3^{128}\gamma^{384} \\ &+ C_2^4C_3^2\gamma^6 + C_2^4C_3^{16}\gamma^{20} + C_2^4C_3^{128}\gamma^{132}C_2^{32}C_3^2\gamma^{34} \\ &+ C_2^{32}C_3^{16}\gamma^{48} + C_2^{32}C_3^{128}\gamma^{160}]\end{aligned}$$

Similarly, one can compute $\Gamma_2(\text{Tr } P(x)\gamma^{2^j})$, where $j \leq 8$ and take the sum over all $0 \leq j \leq 8$ to obtain $\sum_{j\in J}\Gamma_2(\text{Tr } P(x)\gamma^j) = 0$. This results from the choice of the normal basis so that

$$\text{Tr } \gamma^j = 1,\ j \in J,\ \text{Tr } \gamma^{j+k} = 0,\ j,k \in J$$

It has been demonstrated that the binary weight of any codeword in the RS code above is a multiple of 4.

III. Structure of the Code

An examination of the binary version of the RS code reveals nine words whose cyclic components form a (9,6;4) and a (9,3;7) code over $GF(8)$. Thus, the symbol weights are 4, 5, 6, 7, 8, and 9. For weights 4, 5, and 6, one has binary weights of the code 16, 20, and 24. When the (9,3;7) code is nonzero, one has a minimum code weight of 14. But, since the codewords have weights that are multiples of 4, one has a minimum weight of 16.

Now consider C_0 by itself when Tr $C_0 = 0$. This is a binary (72,35;16) code. The encoding here has ceased to be systematic since the condition Tr $C_0 = 0$ is nonsystematic.

The extension parity rule for C_0 has been changed to give the following: For $i = 1$, Tr $P(x)\gamma^i = 1$, and Tr $P(x)\gamma^i = 0$, where $i \neq 1$, an eighth row of weight 8 is adjoined, e.g., 011111111. The argument invoked above [1] can then be used to prove that the minimum code distance is 15 and all odd weight words have weights of form $4m - 1$.

IV. Decoding Binary

Assume at first that Tr $C_0 = 0$. Then the parity sums over all nine binary codes give odd error-pattern information. Thus, with seven errors or less spread out over the nine words, at least two values of $C_3\gamma^j$ must be correct. A trial of eight other values for C_3 will eliminate the cyclic component attached to x^6. Since each of the nine binary codes is now an odd-error-detecting/single-error-correcting code, the parity information is usable to correct single errors when they occur. In the case of seven single binary error patterns in seven different rows, a complete correct decoding emerges. This is the most complex decoding to perform, as it requires eight decoding trials. If the seven or less error patterns are in a smaller set of the nine binary codes, then once C_3 is determined and single errors are corrected, there will be at least six correct coefficients of x.

Undecodable eight-error patterns occur when there are less than three of the nine binary codes that are error-free. Once the x^6 coefficient is obtainable without too many trials, then there must emerge, after single-error correction of the erasure codes, at least six correct coefficients of x. One can of course try 512 values of C_3 and then correct for single errors in what remains and finally decode the remaining (9,6;4) code over $GF(8)$.

Reference

[1] G. Solomon, "Self-Dual (48,24;12) Codes," *The Telecommunications and Data Acquisition Progress Report 42-111, vol. July-September*, pp. 75–79, November 15, 1992.